学位論文

Accuracy of target delineation by positron emission tomography-based auto- segmentation methods after deformable image registration: A phantom study

香川大学大学院医学系研究科

医学専攻

片山　博貴

Technical note

Accuracy of target delineation by positron emission tomography-based auto-segmentation methods after deformable image registration: A phantom study

Hiroki Katayama[a,b,*], Shigeo Takahashi[c], Yukito Maeda[a], Takuya Kobata[a], Akihiro Oishi[a], Yasuhiro Sasakawa[a], Toru Shibata[c]

[a] Department of Clinical Radiology, Kagawa University Hospital, Kagawa, Japan
[b] Graduate School of Medicine, Kagawa University, Kagawa, Japan
[c] Department of Radiation Oncology, Kagawa University Hospital, Kagawa, Japan

ARTICLE INFO

ABSTRACT

Purpose: Diagnostic positron emission tomography and computed tomography (PET/CT) images can be fused to the planning CT images by a deformable image registration (DIR). The aim of this study was to evaluate the standardized uptake value (SUV) and target delineation on deformed PET images.
Methods: We used a cylindrical phantom and removable inserts of four spheres (16–38 mm in diameter) and three ellipsoids with a volume equal to the 38-mm-diameter sphere (S38) in each. S38 was filled with 18F-fluorodeoxyglucose activity, and then PET/CT images were acquired. The contours of S38 were generated using original PET images by PET auto-segmentation (PET-AS) methods of (1) SUV2.5, (2) 40% of maximum SUV (SUV40%max), and (3) gradient-based (GB), and were deformed to the other inserts by DIR. We compared the volumes and the SUV_{max} with the generated contours using the deformed PET images.
Results: The SUV_{max} was slightly decreased by DIR; the mean absolute difference was -0.10 ± 0.04. For SUV2.5 and SUV40%max, the differences in S38 volumes between the original and deformed PET images were less than 5%, regardless of deformation type. For the GB, the contoured volumes obtained from deformed PET images were larger than those of the original PET images for the deformation type of ellipsoids. When the S38 was deformed to the 16-mm-diameter sphere, the maximum volume difference was -22.8%.
Conclusions: Although SUV fluctuations by DIR were negligible, the target delineation on deformed PET images by the GB should be carefully considered owing to the distortion of intensity profiles.

1. Introduction

Positron emission tomography and computed tomography (PET/CT)-guided radiation treatment planning can improve the accuracy of tumor delineation [1] and reduce inter-observer variability [2]. Moreover, the American Association of Physicists in Medicine Task Group 211 (AAPM TG211) [3] has reported a large variety of PET auto-segmentation (PET-AS) methods to assist with definitive a gross tumor volume (GTV) delineation. Fixed threshold algorithms are based on a standardized uptake value (SUV) and the absolute threshold or the percentage of maximum SUV is used in the target delineation [4,5]. Several researchers [6,7] reported that the appropriate threshold values depend on the target size and target-to-background (T/B) ratio. In addition, advanced PET-AS algorithms has been developed and are more

accurate than threshold algorithms [8–10]. Gradient-based method (GB) contours the tumor by the intensity profiles at the target border [11]. However, the quality of segmentation depends on the accuracy of the gradient information [3]. Moreover, machine learning technology has been used to improve the robustness and repeatability of the PET-AS methods in the current study [12,13].

In recent years, diagnostic PET/CT images obtained in a non-treatment position could be incorporated into radiation treatment planning by applying a deformable image registration (DIR) [14–19]. The DIR is performed between the CT components of PET/CT images and planning CT image, and then the PET images are automatically co-registered using the same registration vector as the one used for the CT images [15].

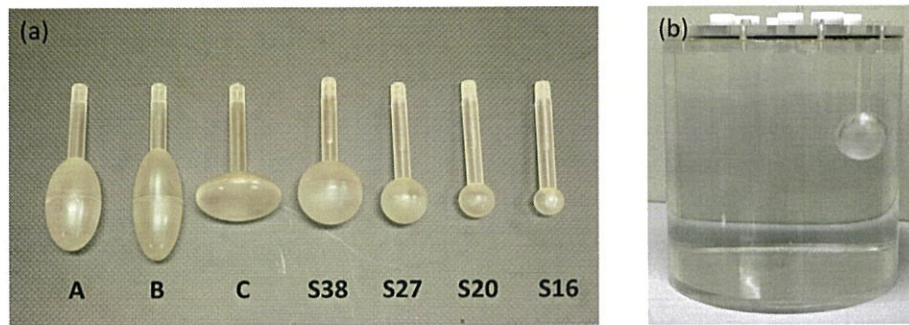However, the use of DIR to combine diagnostic PET/CT images with

Fig. 1. A cylindrical phantom with seven removable inserts was used in this study. (a) Three ellipsoids (A–C) with volumes equal to the S38 and four spheres (S38–S16) with diameters of 38 mm, 27 mm, 20 mm, and 16 mm, respectively. (b) Insertion of S38 attached to the cylindrical phantom.

planning CT images has not been consistently validated for radiation treatment planning, specifically, deformed PET images have not been used to assess a SUV and target delineation [20–23]. In a previous study, only the contours generated by the PET-AS using original PET images were deformed because the voxel values on PET images can change due to the deformation and interpolation effects [20,24].

We simulated target deformations using phantoms with various shapes and volumes, and investigated the difference in maximum SUV (SUV_{max}) of the FDG-avid region before and after DIR. Moreover, we compared the contour generated by the PET-AS method using the deformed PET images with the generated contour before DIR, validated the accuracy and consistency of target delineation in various deformation types and slice thicknesses of PET/CT images.

2. Materials and methods

2.1. Phantom object

Fig. 1 shows a cylindrical phantom and seven types of removable inserts (Kobe Itoi Factory Inc, Japan) that were used in this study. The removable inserts were composed of three ellipsoids (type A–C) and four spheres of different diameters (S38–S16). The ellipsoid inserts were custom-made for this study, corresponding to the volume of a 38-mm-diameter sphere (28.7 mL). The diameters of type A in the lateral, vertical, and longitudinal directions were 31, 31, and 57 mm, respectively, whereas those of type B and C were 28, 28, and 70 mm and 50, 50, and 22 mm, respectively. The diameters of types S38, S27, S20, and S16 were 38, 27, 20, and 16 mm, respectively.

2.2. PET/CT acquisition

The insert of S38 was defined as original PET/CT images and deformed to the other inserts. The cylindrical phantom was filled with an ^{18}F-fluorodeoxyglucose (FDG) background with the concentration of 860 Bq/mL. The insert of S38 was attached to the cylindrical phantom and was filled with a mixture of FDG and iodine contrast agent to achieve T/B ratios of 5, 10, and 15. The iodine contrast agent was used to improve the accuracy of CT–CT registration.

We used a Biograph mCT 64-slice scanner (Siemens Medical Solutions, Erlangen, Germany) with a transverse field of view (FOV) of 70 cm, and an axial FOV of 21.6 cm. PET data were acquired in the three-dimensional list mode for 10 min.

PET scan parameters were bed position, 1; scan time, 600 s; FOV, 700 mm; and matrix size, 256 × 256 with 3.18 mm pixels. Reconstruction parameters were 3D-OSEM with point-spread function and time-of-flight (2 iterations and 21 subsets), and a Gaussian filter with a full width at half maximum of 5 mm.

CT scan parameters included the following: scanning mode, helical; tube voltage, 120 kV; tube current, 50 mA; rotation time, 0.5 s; helical pitch, 0.8; FOV, 500 mm; and matrix size, 512 × 512 with 0.98 mm

pixels. The slice thickness of PET/CT acquired was 2, 5, and 7 mm.

2.3. Planning CT acquisition

The inserts of A–C and S27–S16 were filled only with iodine contrast agent on the assumption that the target was deformed. Each insert was attached one by one to the cylindrical phantom, and planning CT images were acquired on Aquilion LB (Canon Medical Systems Corp., Tokyo, Japan) with a slice thickness of 2 mm. Scan parameters included the following: scanning mode, helical; tube voltage, 120 kV; tube current, auto mA (max 420 mA); rotation time, 0.5 s; helical pitch, 0.94; FOV, 500 mm; and matrix size, 512 × 512 with 0.98 mm pixels.

2.4. Deformable image registration

Fig. 2 shows the flow chart of the creation of deformed PET/CT images and the experiments. The PET/CT and planning CT image data were transferred to MIMmaestro Ver. 6.4 (MIM Software Inc, Cleveland, OH, USA). The insert of S38 was deformed to the other inserts on planning CT image by DIR. First, rigid image registration was performed between CT components of the PET/CT and planning CT images, and the DIR was subsequently performed over the entire image. The DIR algorithm with MIM is an intensity-based free-form deformation algorithm with essentially limitless degrees of freedom, which was
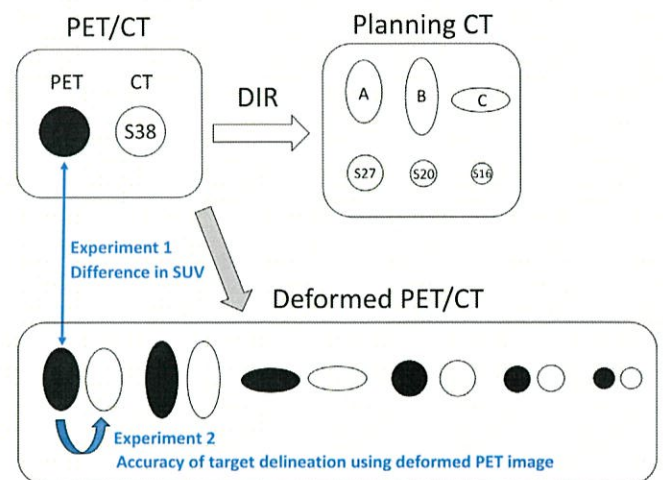

Fig. 2. Flow chart of the creation of deformed PET/CT images and the experiments. The DIR was performed between the CT images of S38 and the CT images of the other inserts. Next, the deformed PET images were created by applying the deformation vector field of CT–CT registration to the PET images. Experiment 1 measured the difference in SUV between the original PET images and deformed PET images (type A-C and S27-S16). Experiment 2 compared the accuracy of target delineation using the deformed PET image (type A-C and S27-S16) with the contour generated before DIR.
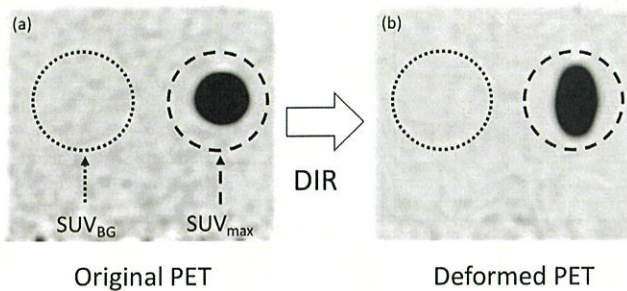
Fig. 3. (a) The PET images of the S38 and (b) deformed PET images applied to the insert of A. The SUV_{max} and SUV_{BG} were obtained from the VOI of the dashed line and dotted line.

previously evaluated [20,25]. Finally, deformed PET images were created by applying the deformation vector field (DVF) of CT–CT registration to PET images because the PET/CT image pairs are in the same DICOM frame of reference [22]. The interpolation process of PET images was performed only at the final step.

2.5. SUV measurement of original PET and deformed PET images (Experiment 1)

Fig. 3 shows S38 on the original PET and of the deformed PET images. The SUV_{max} was established by drawing the volume of interest (VOI, 8 cm in diameter) encompassing the target. To assess the SUV in the background regions, the VOI was placed into the relevant background structure at a safe distance from the target, and the mean SUV was defined as SUV_{BG}. We measured the SUV_{max} and SUV_{BG} on the deformed PET images of the A-C and S27-S16 and compared them with the original PET images. The measurements were performed on the PET/CT images for T/B ratios of 5, 10, and 15 with a slice thickness of 2 mm.

2.6. Target delineation by PET-AS methods using original and deformed PET images (Experiment 2)

The validation of PET-AS methods in this study was performed on fixed threshold values of SUV2.5 (SUV2.5), threshold of 40% of maximum SUV_{max}(SUV40%max), and the gradient-based (GB) method using PET Edge by MIM [5,9]. The threshold values of segmentation methods were determined to delineate S38 under the conditions of T/B ratio 5 because the appropriate threshold value depended on the target size and T/B ratio [6,26,27].

Fig. 4 shows the target delineation procedures generated by the PET-AS methods using original PET and deformed PET images.

2.6.1. Target delineation using original PET images

The contours of S38 were generated by the PET-AS methods using the original PET images and were then deformed to other inserts by DIR. We assessed the delineation accuracy of all segmentations performed by comparing these contours with the contour obtained for CT-derived ground truth (CT_GT). The CT_GT was determined by drawing a sphere of the same diameter as S38 and positioning it on associated CT images to match the sphere wall. The contours of SUV2.5, SUV40%max, GB, and CT_GT on the deformed CT images were defined as $SUV2.5_{pre}$, $SUV40\%max_{pre}$, GB_{pre}, and CT_GT_{pre}, respectively.

2.6.2. Target delineation using deformed PET images

The S38 on the deformed CT images applied to each insert were contoured by the PET-AS methods using deformed PET images. The delineated contours were defined as $SUV2.5_{post}$, $SUV40\%max_{post}$, and GB_{post}, respectively. The volume differences in contours were calculated for each PET-AS method as follows:
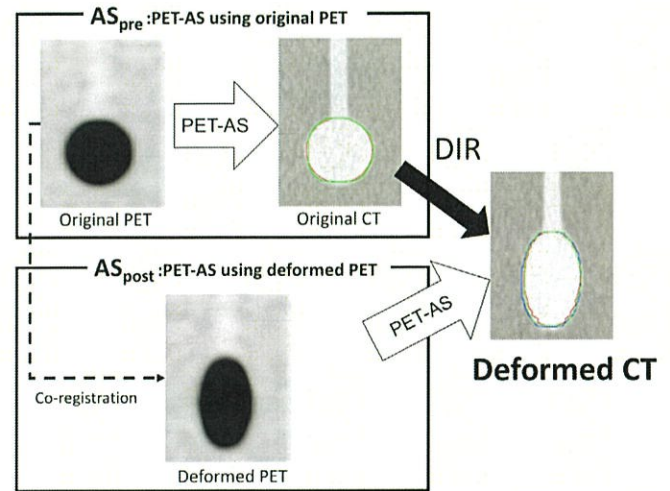

Fig. 4. The procedures of target delineation generated by the PET-AS methods using original PET and deformed PET images. The red and green lines on the original CT images are contour of the S38 generated by CT-based and a PET-AS method (GB), and those on the deformed CT images are applying the DIR to insert of A. The blue line on the deformed CT images is contour of the S38 generated by the GB using deformed PET images. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$Volume\ difference(\%) = \frac{(AS_{post} - AS_{pre})}{AS_{pre}} \times 100 \qquad (1)$$

where AS indicates the volume of contour generated by each PET-AS method.

We performed target delineation on the PET/CT image with the slice thickness of 2, 5, and 7 mm, and investigated the difference in volume resulting from slice thickness.

Next, we compared the accuracy of target delineation between the original and deformed PET images. The AS_{pre} and AS_{post} were compared with CT_GT_{pre} in terms of volume and geometrical overlap using the dice similarity coefficient (DSC) [28]. The DSC is a spatial overlap index that assigns values in the [0, 1] range with 0 corresponding to no overlap and 1 corresponding to a complete overlap [29]. For each slice thickness of PET/CT images, we calculated the mean value of DSC for deformed inserts and compared them between AS_{pre} and AS_{post}. Moreover, we evaluated the concordance of DSC values between AS_{pre} and AS_{post} using an Intra Class Correlation (ICC) test in each PET-AS method.

2.7. Statistical analysis

We performed a series of experiments from PET/CT and planning CT image acquisition to DIR four times independently on different days for the evaluation of repeatability and indicated all values as mean ± standard error (SE). Statistical analysis was performed using JMP® 13 (SAS Institute Inc., Cary, NC, USA). A Wilcoxon signed-rank test was used to determine whether significant differences in DSC existed between AS_{pre} and AS_{post} for each slice thickness of PET/CT images. A P-value of < 0.05 was considered significant for all statistical tests. We tested the concordance of DSC values between AS_{pre} and AS_{post} with a two-way random model ICC with absolute agreement using the Statistical Package for Social Sciences (SPSS, Ver. 22, SPSS Inc., Chicago, IL, USA).
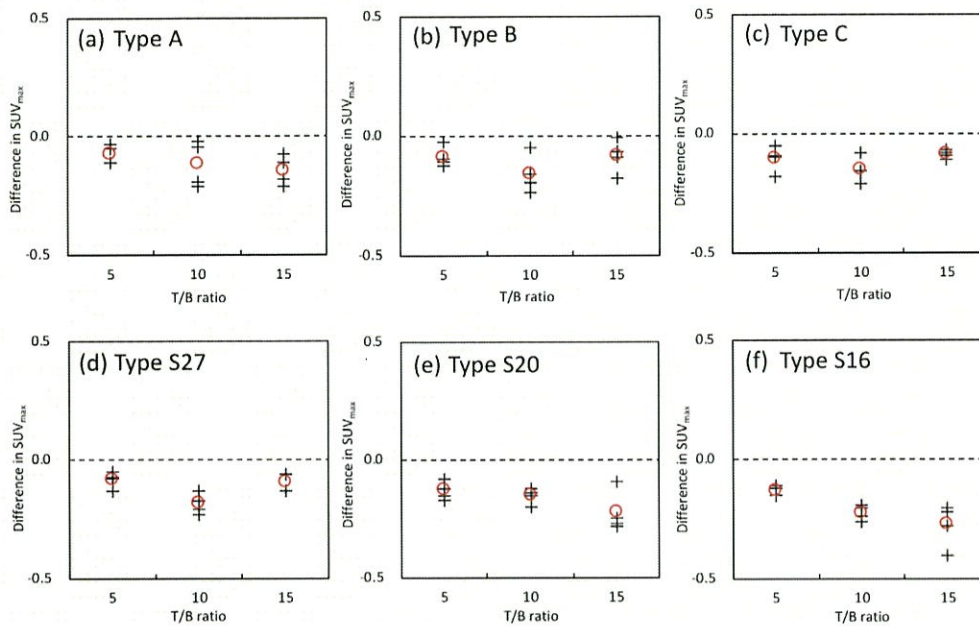
Fig. 5. Difference in SUV_{max} of the S38 between original PET and deformed PET images applied to the inserts of (a-c) A–C and (d-f) S27–S16 for T/B ratios of 5, 10, and 15. The cross markers indicated the measurement data, and the red open circle indicated the mean value of four times. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Difference in SUV_{max} between before and after DIR

Fig. 5 (a-c) shows the difference in SUV_{max} of S38 between the original PET and deformed PET images of A–C. For T/B ratio of 5, the means of absolute differences were -0.08 (range, -0.05 to -0.11), -0.09 (range, -0.03 to -0.13), -0.11 (range, -0.05 to -0.18) for deformation types of A, B, and C, respectively. The SUV_{max} of the deformed PET images slightly decreased from the original PET images for all T/B ratios.

Fig. 5 (d-f) shows the difference in SUV_{max} of S38 between the original PET and deformed PET images of S27-S16. For T/B ratio of 5, the means of absolute differences were -0.08 (range, -0.05 to -0.13), -0.13 (range, -0.08 to -0.17), -0.13 (range, -0.11 to -0.15) for deformation types of S27, S20, and S16 respectively. When S38 was deformed to the S16, the SUV_{max} was stronger decreased for all T/B ratios. The summarized data is presented in supplementary Table S1.

The voxel intensities in the deformed PET images were determined by interpolation of the original PET intensities [30]. Thus, the SUV_{max} of the deformed PET images was always lower than that of the original PET images.

The fluctuations of SUV in the background regions were less than 0.01 for all T/B ratios, and changes in voxel intensities following DIR application to the PET images were negligible.

3.2. Difference in contour volume between AS_{pre} and AS_{post}

We compared the contour volumes between the two delineation procedures. Fig. 6 shows the percentage difference in contour volumes between AS_{pre} and AS_{post} for the deformation types of A–C.

For the SUV2.5 and SUV40%max methods, the difference in contour volumes between AS_{pre} and AS_{post} was less than 5% in each slice thickness of PET/CT images. The maximum volume difference (%) was 1.9% $\pm$ 0.9% in the SUV2.5 and 4.0% $\pm$ 1.4% in the SUV40%max at the deformation types of B and a PET/CT slice thickness of 7 mm. For the GB method, the contour volume of AS_{post} was found to be larger than that of AS_{pre}. For the PET/CT slice thickness of 2 mm, the volume differences (%) of A, B, and C were 6.6% $\pm$ 1.0%, 9.6% $\pm$ 0.8%, and 9.8% $\pm$ 0.6%, respectively. The difference in contour volumes between

AS_{pre} and AS_{post} increased for the PET/CT slice thicknesses of 5 and 7 mm, which were found at the deformation type of B. The summarized data is presented in supplementary Table S2.

Fig. 7 shows the percentage difference in contour volume between AS_{pre} and AS_{post} for the deformation types of S27–S16.

For the SUV2.5 and SUV40%max methods, the contour volumes of AS_{post} corresponded to the AS_{pre} in each slice thickness of PET/CT images. The maximum volume difference was 1.8% $\pm$ 0.8% in SUV2.5 and 5.0% $\pm$ 0.9% in SUV40%max at the deformation of S27 and a slice thickness of 7 mm. The contour volume of GB_{post} increased for the deformation type of S20 and decreased for the deformation type of S16 compared with that of GB_{pre}. The summarized data is presented in supplementary Table S3.

3.3. Accuracy of PET-AS methods using original PET and deformed PET images

We evaluated the accuracy of target delineation by the PET-AS methods using deformed PET images for each slice thickness of the PET/CT images.

Fig. 8 shows the DSC values obtained from between the CT_GT_{pre} and the contours of AS_{pre} and AS_{post} for each slice thickness of PET/CT image. For the SUV2.5 and SUV40%max methods, the DSC values were not significantly different between the AS_{pre} and AS_{post} at a slice thickness of 2 mm (P = 0.0521, P = 0.9561). For the GB method, the DSC values of AS_{post} were 0.90 $\pm$ 0.02, 0.88 $\pm$ 0.02, and 0.88 $\pm$ 0.02 in the slice thicknesses of 2, 5, and 7 mm, respectively, which were lower than those of AS_{pre} (P < 0.001). The summarized data is presented in supplementary Table S4.

Table 1 summarizes the ICC values obtained from the comparison of DSC between AS_{pre} and AS_{post} in each PET-AS method. At a slice thickness of 2 mm, the DSC values of the contours indicated the high concordance between AS_{pre} and AS_{post} when generated by the SUV2.5 and SUV40%max methods (ICC > 0.8), but lower concordance with GB method (ICC = 0.17). Moreover, the ICC values were lower in association with the slice thicknesses of PET/CT images in all methods.

4. Discussion

We evaluated the voxel intensities of the deformed PET images and the accuracy and consistency of the PET-AS methods. The SUV_{max} was
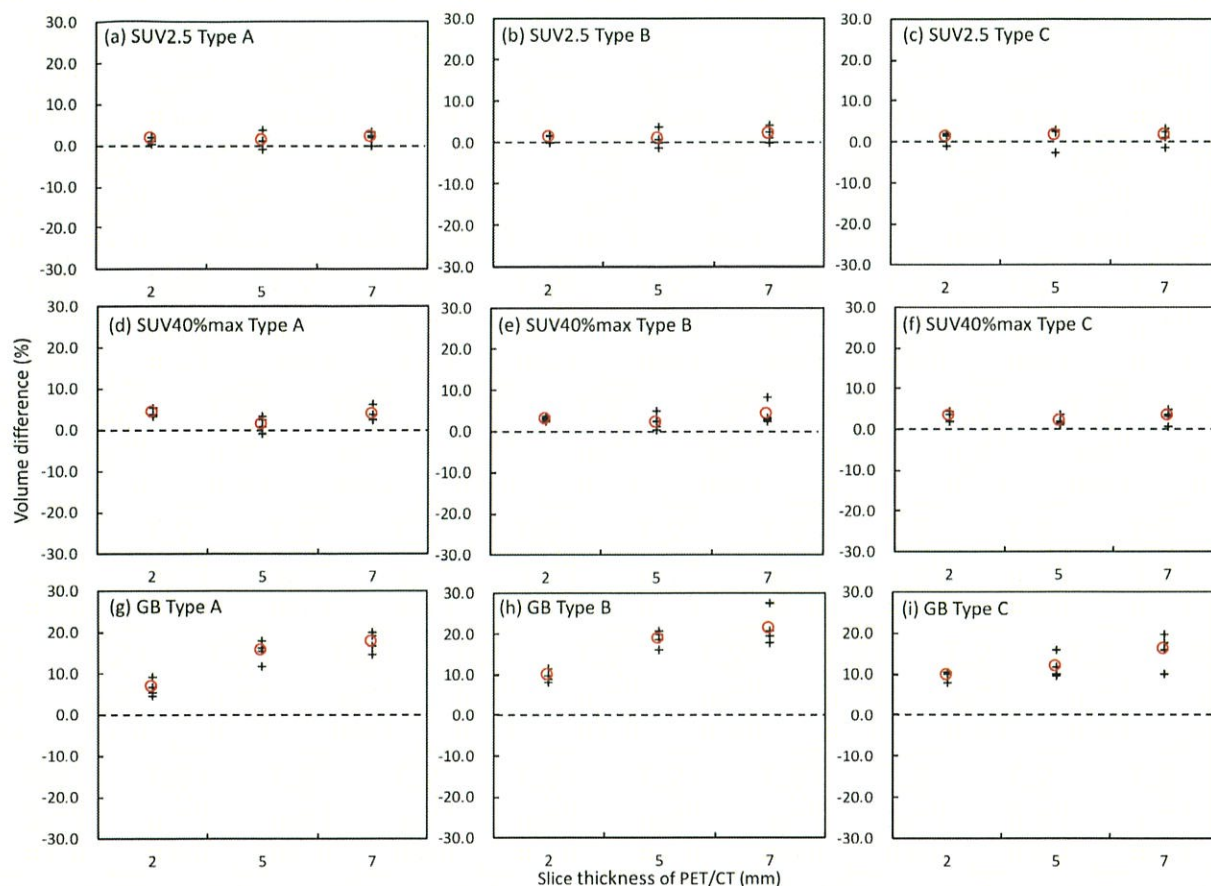
Fig. 6. Volume difference (%) in contour between AS_{pre} and AS_{post} generated by the PET-AS methods of (a-c) SUV2.5, (d-f) SUVmax40%, and (g-i) GB for deformation type of A–C. The horizontal axis indicates the slice thickness of PET/CT images. The cross markers indicated the measurement data, and the red open circle indicated the mean value of four times. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

slightly decreased after applying DIR to the PET images, and the decreased the SUV_{max} was found to be associated with the deformation type.

We considered that the voxel intensities of the deformed PET images were determined by the interpolation of the original PET intensities. The PET voxel was mapped to a new position based on the DVF used on CT–CT registration, resulting in new PET/CT images with deformed registration with the planning CT [16]. Thus, the voxel size of the deformed PET images corresponded to the planning CT images. The SUV_{max} is the highest intensity value in the original PET image, and interpolation is a kind of "average" between nearby voxel intensities. Consequently, the interpolated values will always be lower than the original SUV_{max}. If the original PET insert is deformed to smaller volumes, the SUV_{max} was stronger decreased. Because of the reduced number of voxels in the smaller inserts, fewer interpolated PET intensities will be calculated, so the chance to obtain a high SUV_{max} will decrease. For background regions, the SUV_{BG} of the deformed PET images were not always lower than that of the original PET images. However, the differences in SUV_{BG} between the original PET and deformed PET images were negligible.

Although diagnostic PET/CT images can be fused to the planning CT images by DIR, utilization of the deformed PET images has not yet been clarified. Here, we compared the contours generated by the PET-AS methods using the original and deformed PET images. We considered that the intensity profile could be the cause for delineation volume difference. The GB method was processed using denoising and deblurring from intensity profiles of the PET images [11], and the points where the gradient was the largest in magnitude corresponded to the target contours [3]. Therefore, the difference in the contour volume was induced due to a slight distortion of the target border profiles. When the target volume was decreased by DIR, the delineating volume was over or underestimated for significant distortion of the target border.

When the slice thicknesses of the PET/CT images were 5 and 7 mm, the contour of AS_{pre} was deformed to an irregular contour in the slice direction by DIR. On the other hand, the contour of the AS_{post} was delineated using deformed PET images with a slice thickness of 2 mm, especially the contour of GB_{post} was delineated around the GB_{pre} owing to the distortion of profiles. Thus, the volume difference was associated with the number of slices for delineating target shapes such as the deformation types of A and B.

We investigated the accuracy of PET-AS using the deformed PET images by quantitatively comparing target volumes defined by auto segmentation using the DSC.

For target delineation using the original PET image, threshold segmentation methods such as SUV2.5 (absolute threshold) and SUV40% max (percentage of maximum) are widely used for their simplicity and ease [3–5]. The GB method is a robust segmentation technique for different PET cameras and imaging conditions using image gradient information [9,31], which is always in good agreement regardless of the T/B ratio. Berthon et al. [26] reported that segmentation with fixed thresholding methods and GB methods were in good agreement at a higher T/B ratio (> 5). For target delineation using the deformed PET image, the contours generated by the threshold segmentation methods was close to the contour generated before DIR. The deformed PET images were created using the DVF of CT-CT registration, and intensity fluctuations of the PET voxels by DIR were negligible. Accordingly, the appropriate threshold value for target delineation corresponded to that before DIR regardless of the deformation type. In contrast, the contour
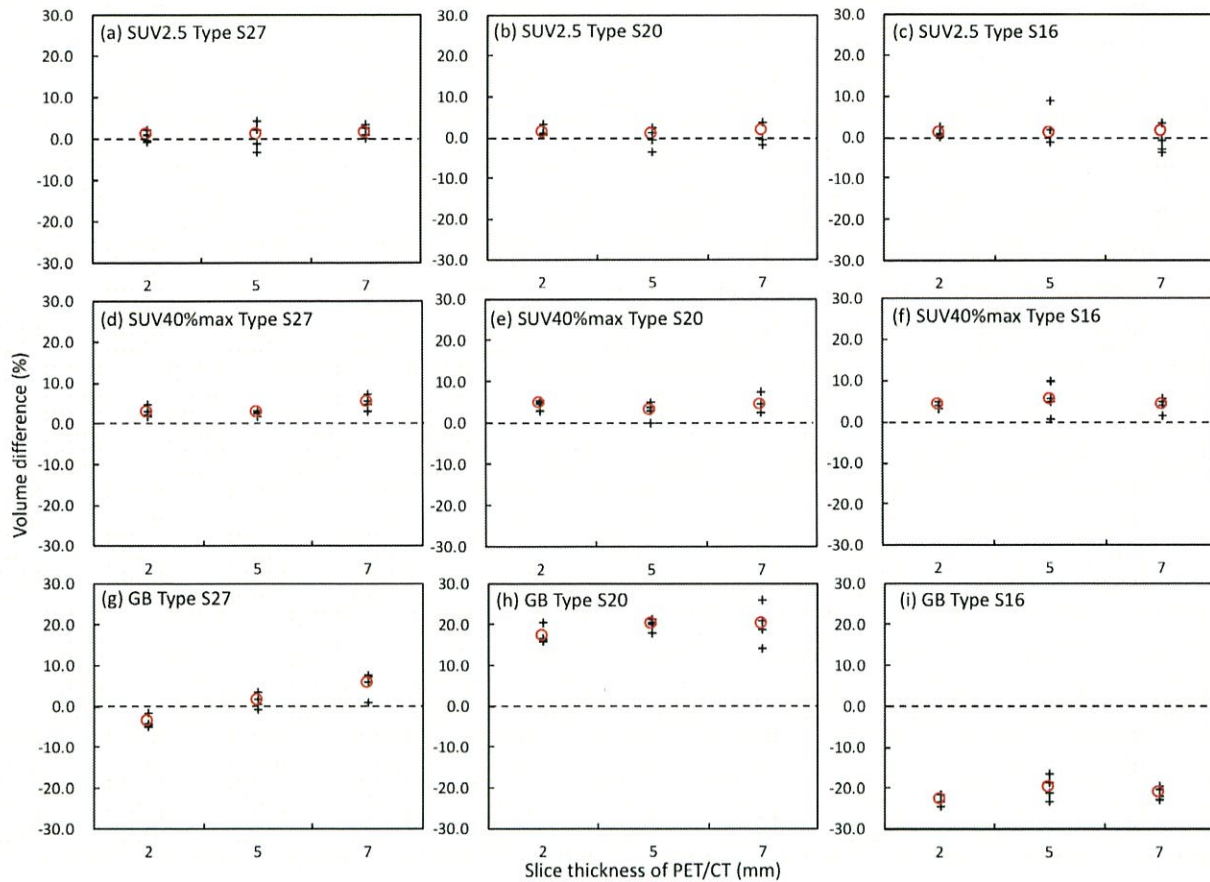
Fig. 7. Volume difference (%) in contour between AS_{pre} and AS_{post} generated by the PET-AS methods of (a-c) SUV2.5, (d-f) SUVmax40%, and (g-i) GB for deformation types of S27–S16. The horizontal axis indicates the slice thickness of the PET/CT images. The cross markers indicated the measurement data, and the red open circle indicated the mean value of four times. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with the GB method was delineated outside the CT_GT owing to distortion of the intensity profile, and the value of DSC was lower in association with the slice thickness of PET/CT images. This study indicated that the segmentation by GB should be performed on the original PET images, and that contour deformed afterwards.

In previous studies exploring the use of diagnostic PET/CT images in treatment planning, the PET-segmentation strategies have never been clarified. Our results indicated that the GTV delineation using PET-AS with deformed PET images affects tumor borders, depending on the segmentation algorithm and slice thickness of PET/CT images. These differences may result in increased risk to the adjacent organ at risk (OAR) or decreased dose to the target. In particular, the contour generated by the GB method was strongly affected by the deformation of PET images, which decreased the accuracy of GTV delineation. Therefore, without information regarding the impact of DIR on PET image, the difference in contour volume by the delineation procedure will be missed.

Our study has some limitations. First, to clarify the relationship between the target border on CT and SUV, we assumed that a target had adequate volume and homogeneous activity using the phantom. Namely, the fluctuations of PET intensities in the target and background regions were determined by PET noise. Thus, validation of auto-segmentation in realistic clinical lesions where PET uptake is in-homogeneous is needed for clinical use. Future work will focus on noting the change in SUV on clinical images. Second, changes in the SUV and the generated contour using the deformed PET images may be dependent on the interpolator within the DIR software. The voxel intensities on the deformed PET images were interpolated by an intrinsic method in the MIM maestro. Thus, the SUV_{max} in the target after DIR

might have different intensity values and, depending on the inter-polation methods, the voxel intensities of the tumor border might be different values. Third, this study assumed the tumor was unaffected by the respiratory motion and was designed to focus on the change in SUV by DIR using the static phantoms. For lung and liver tumors, the volume defined by PET-segmentation represents the internal target volume because the PET images are blurred by the respiratory motion [32]. Hence, we propose that applying DIR to the PET images with respiratory motion should be investigated in further independent studies. Fourth, the accuracy of target delineation is mostly dependent on CT-CT registration. The anatomical correspondence of the tumor and organ at risk is necessary between planning CT and CT components of the PET/CT after DIR [23]. Hence, careful review of the delineated tumor is mandatory [24]. Fifth, PET-based target delineation must consider the possible misalignment between the CT and PET images [33]; moreover, GTV delineation using thick PET images may include more than the metabolically active part of the tumor caused by the partial volume effect [34].

5. Conclusions

We evaluated the accuracy and consistency of the PET-AS methods using deformed PET images. The SUV fluctuations of PET images by DIR were negligible. The threshold segmentation methods using the SUV_{max} and the background intensities might be used in the target delineation of the deformed PET images, while the GB methods using image gradient information should be carefully considered owing to distortion of intensity profiles. In particular, the slices of diagnostic PET/CT images were thicker than those of the planning CT images in
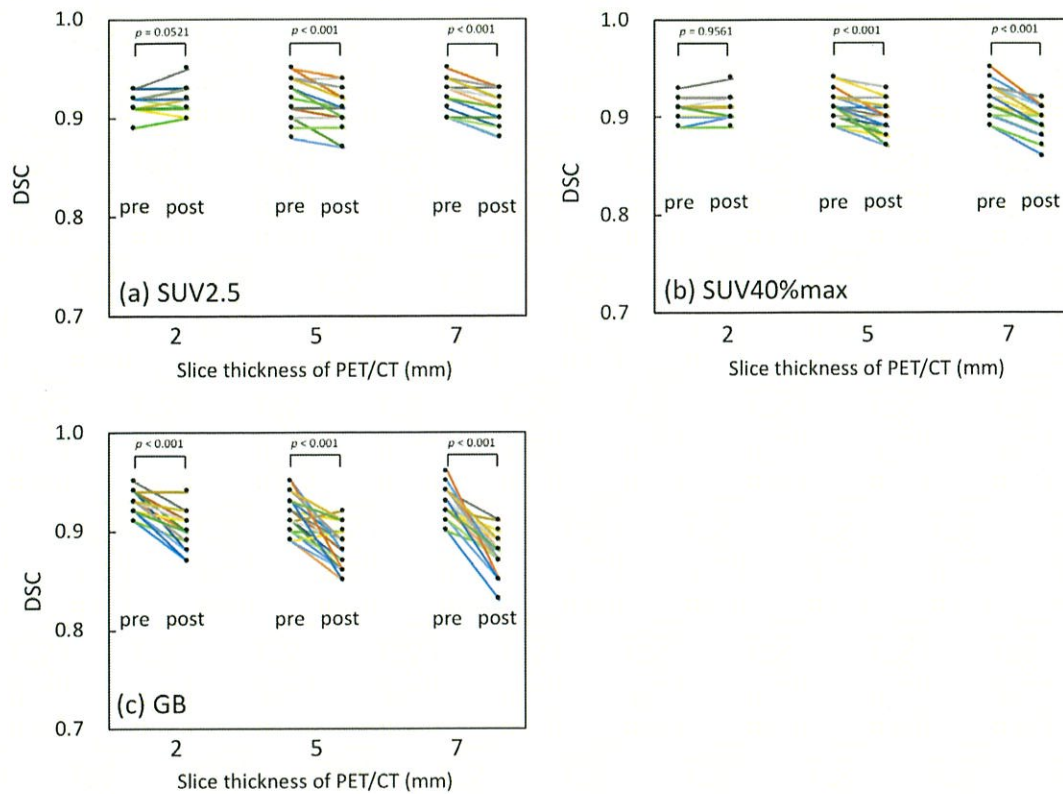
Fig. 8. DSC values obtained from the contour of AS_{pre} and AS_{post} generated by the PET-AS methods of (a) SUV2.5, (b) SUVmax40%, and (c) GB for each slice thickness of PET/CT image. The corresponding DSC values of AS_{pre} (left point) and AS_{post} (right point) were connected using a line.

Table 1
ICC values and corresponding 95% confidence intervals of the DSC values between AS_{pre} and AS_{post} in each PET-AS method.

PET-AS method	Slice thickness of PET/CT (mm)		
	2	5	7
SUV2.5	0.80 (0.51 to 0.91)	0.81 (0.31 to 0.93)	0.69 (−0.70 to 0.92)
SUV40%max	0.87 (0.73 to 0.94)	0.65 (0.05 to 0.87)	0.49 (−0.87 to 0.82)
GB	0.17 (−0.08 to 0.51)	0.07 (−0.09 to 0.31)	0.02 (−0.05 to 0.15)

most cases. Thus, the segmentation by the GB method using the deformed PET images decreased the accuracy of GTV delineation and may be the cause of increased risk to the OAR or decreased dose to the target.

Funding

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmp.2020.07.015.

References

[1] Paulino AC, Koshy M, Howell R, Schuster D, Davis LW. Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2005;61:1385–92. https://doi.org/10.1016/j.ijrobp.2004.08.037.

[2] Karki K, Saraiya S, Hugo GD, Mukhopadhyay N, Jan N, Schuster J, et al. Variabilities of MRI-, CT- and PET-CT-based tumor and lymph node delineations for lung cancer radiotherapy planning. Int J Radiat Oncol Biol Phys 2017;99:80–9. https://doi.org/10.1016/j.ijrobp.2017.05.002.

[3] Hatt M, Lee JA, Schmidtlein CR, Lu W, Jeraj R. Classification and evaluation strategies of auto-segmentation approaches for PET : Report of AAPM task group No.211. Med Phys 2017;44.

[4] Hong R, Halama J, Bova D, Sethi A, Emami B. Correlation of PET standard uptake value and CT window-level thresholds for target delineation in CT-based radiation treatment planning. Int J Radiat Oncol Biol Phys 2007;67:720–6. https://doi.org/10.1016/j.ijrobp.2006.09.039.

[5] Dalah E, Moraru I, Paulson E, Erickson B, Li XA. Variability of target and normal structure delineation using multimodality imaging for radiation therapy of pancreatic cancer. Int J Radiat Oncol Biol Phys 2014;89:633–40. https://doi.org/10.1016/j.ijrobp.2014.02.035.

[6] Biehl KJ, Kong F-M, Dehdashti F, Jin J-Y, Mutic S, El Naqa I, et al. 18F-FDG PET definition of gross tumor volume for radiotherapy of non-small cell lung cancer: is a single standardized uptake value threshold approach appropriate? J Nucl Med 2006;47:1808–12. 47/11/1808 [pii].

[7] Ford EC, Kinahan PE, Hanlon L, Alessio A, Rajendran J, Schwartz DL, et al. Tumor delineation using PET in head and neck cancers: threshold contouring and lesion volumes. Med Phys 2006;33:4280–8. https://doi.org/10.1118/1.2361076.

[8] Day E, Betler J, Parda D, Reitz B, Kirichenko A, Mohammadi S, et al. A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients. Med Phys 2009;36:4349–58. https://doi.org/10.1118/1.3213099.

[9] Werner-Wasik M, Nelson AD, Choi W, Arai Y, Faulhaber PF, Kang P, et al. What is the best way to contour lung tumors on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. Int J Radiat Oncol Biol

Phys 2012;82:1164–71. https://doi.org/10.1016/j.ijrobp.2010.12.055.

[10] Berthon B, Evans M, Marshall C, Palaniappan N, Cole N, Jayaprakasam V, et al. Head and neck target delineation using a novel PET automatic segmentation algorithm. Radiother Oncol 2017;122:242–7. https://doi.org/10.1016/j.radonc.2016.12.008.

[11] Geets X, Lee JA, Bol A, Lonneux M, Grégoire V. A gradient-based method for segmenting FDG-PET images: Methodology and validation. Eur J Nucl Med Mol Imaging 2007;34:1427–38. https://doi.org/10.1007/s00259-006-0363-4.

[12] Parkinson C, Evans M, Guerrero-Urbano T, Michaelidou A, Pike L, Barrington S, et al. Machine-learned target volume delineation of 18F-FDG PET images after one cycle of induction chemotherapy. Phys Medica 2019;61:85–93. https://doi.org/10.1016/j.ejmp.2019.04.020.

[13] Comelli A, Bignardi S, Stefano A, Russo G, Sabini MG, Ippolito M, et al. Development of a new fully three-dimensional methodology for tumours delineation in functional images. Comput Biol Med 2020;120:103701https://doi.org/10.1016/j.compbiomed.2020.103701.

[14] Ward G, Ramasamy S, Sykes JR, Prestwich R, Chowdhury F, Scarsbrook A, et al. Superiority of deformable image co-registration in the integration of diagnostic positron emission tomography-computed tomography to the radiotherapy treatment planning pathway for oesophageal carcinoma. Clin Oncol 2015;28:655–62. https://doi.org/10.1016/j.clon.2016.05.009.

[15] Thomas L, Lapa C, Bundschuh RA, Polat B, Sonke JJ, Guckenberger M. Tumour delineation in oesophageal cancer – A prospective study of delineation in PET and CT with and without endoscopically placed clip markers. Radiother Oncol 2015;116:269–75. https://doi.org/10.1016/j.radonc.2015.07.007.

[16] Fortin D, Basran PS, Berrang T, Peterson D, Wai ES. Deformable versus rigid registration of PET/CT images for radiation treatment planning of head and neck and lung cancer patients: a retrospective dosimetric comparison. Radiat Oncol 2014;9:50. https://doi.org/10.1186/1748-717X-9-50.

[17] Kovalchuk N, Jalisi S, Subramaniam RM, Truong MT. Deformable registration of preoperative PET/CT with postoperative radiation therapy planning CT in head and neck cancer. Radiographics 2012;32:1329–41. https://doi.org/10.1148/rg.325125008.

[18] Cho O, Chun M, Oh Y-T, Kim M-H, Park H-J, Heo J-S, et al. Can initial diagnostic PET-CT aid to localize tumor bed in breast cancer radiotherapy: feasibility study using deformable image registration. Radiat Oncol 2013;8:163. https://doi.org/10.1186/1748-717X-8-163.

[19] Fontanilla HP, Klopp AH, Lindberg ME, Jhingran A, Kelly P, Takiar V, et al. Anatomic distribution of [18F] fluorodeoxyglucose-avid lymph nodes in patients with cervical cancer. Pract Radiat Oncol 2013;3:45–53. https://doi.org/10.1016/j.prro.2012.02.003.

[20] Guo Y, Li J, Zhang P, Shao Q, Xu M, Li Y. Comparative evaluation of target volumes defined by deformable and rigid registration of diagnostic PET/CT to planning CT in primary esophageal cancer. Medicine (Baltimore) 2017;96:e5528https://doi.org/10.1097/MD.0000000000005528.

[21] Nestle U, Weber W, Hentschel M, Grosu A-L. Biological imaging in radiation therapy: role of positron emission tomography. Phys Med Biol 2009;54:R1–25. https://doi.org/10.1088/0031-9155/54/1/R01.

[22] Thorwarth D, Beyer T, Boellaard R, de Ruysscher D, Grgic A, Lee JA, et al. Integration of FDG-PET/CT into external beam radiation therapy planning: technical aspects and recommendations on methodological approaches. Nuklearmedizin 2012;51:140–53. https://doi.org/10.3413/Nukmed-0455-11-12.

[23] Konert T, Vogel W, MacManus MP, Nestle U, Belderbos J, Grégoire V, et al. PET/CT imaging for target volume delineation in curative intent radiotherapy of non-small cell lung cancer: IAEA consensus report 2014. Radiother Oncol 2015;116:27–34. https://doi.org/10.1016/j.radonc.2015.03.014.

[24] Spijkerman J, Fontanarosa D, Das M, van Elmpt W. Validation of nonrigid registration in pretreatment and follow-up PET/CT scans for quantification of tumor residue in lung cancer patients. J Appl Clin Med Phys 2014;15:240–50. https://doi.org/10.1120/jacmp.v15i4.4847.

[25] Pukala J, Johnson PB, Shah AP, Langen KM, Bova FJ, Staton RJ, et al. Benchmarking of five commercial deformable image registration algorithms for head and neck patients. J Appl Clin Med Phys 2016;17:25–40. https://doi.org/10.1120/JACMP.V17I3.5735.

[26] Berthon B, Marshall C, Evans M, Spezi E. Evaluation of advanced automatic PET segmentation methods using nonspherical thin-wall inserts. Med Phys 2014;41:022502https://doi.org/10.1118/1.4863480.

[27] Toya R, Murakami R, Tashiro K, Yoshida M, Sakamoto F, Kawanaka K, et al. FDG-PET/CT-based Gross Tumor Volume Contouring for Radiation Therapy Planning: An Experimental Phantom Study. J Radiat Res 2012;53:338–41. https://doi.org/10.1269/jrr.10183.

[28] Zou KH, Warfield SK, Bharatha A, Tempany CMC, Kaus MR, Haker SJ, et al. Statistical validation of image segmentation quality based on a spatial overlap index. Acad Radiol 2004;11:178–89. https://doi.org/10.1016/S1076-6332(03)00671-8.

[29] Loi G, Fusella M, Lanzi E, Cagni E, Garibaldi C, Iacoviello G, et al. Performance of commercially available deformable image registration platforms for contour propagation using patient-based computational phantoms: a multi-institutional study. Med Phys 2018;45:748–57. https://doi.org/10.1002/mp.12737.

[30] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132: Report. Med Phys 2017;44:e43–76. https://doi.org/10.1002/mp.12256.

[31] Wanet M, Lee JA, Weynand B, De Bast M, Poncelet A, Lacroix V, et al. Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: a comparison with threshold-based approaches, CT and surgical specimens. Radiother Oncol 2011;98:117–25. https://doi.org/10.1016/j.radonc.2010.10.006.

[32] Riegel AC, Kara Bucci M, Mawlawi OR, Ahmad M, Luo D, Chandler A, et al. Defining internal target volume using positron emission tomography for radiation therapy planning of moving lung tumors. J Appl Clin Med Phys 2014;15:279–89. https://doi.org/10.1120/jacmp.v15i1.4600.

[33] Beyer T, Tellmann L, Nickel I, Pietrzyk U. On the use of positioning aids to reduce misregistration in the head and neck in whole-body PET/CT studies. J Nucl Med 2005;46:596–602.

[34] Soret M, Bacharach SL, Buvat I. Partial-Volume Effect in PET Tumor Imaging. J Nucl Med 2007;48:932–45. https://doi.org/10.2967/jnumed.106.035774.